

SPLICING ERROR IN *GATA1* AFFECTS ERYTHROPOIESIS IN THE *XPNA* MOUSE  
(X-LINKED PRE- AND NEONATAL ANEMIA) WITH SUGGESTION OF A NOVEL  
COMPENSATORY ERYTHROID TRANSCRIPTION FACTOR

BY

KYLE LAMAR MILLER

BS, FRIENDS UNIVERSITY, 2010

Submitted to the graduate degree program in Clinical Research and the Graduate Faculty of the  
University of Kansas in partial fulfillment of the requirements for the degree of Master of  
Science.

---

Chairperson Theresa I. Shireman, PhD, RPh

---

Robert A. White, PhD

---

Douglas C. Bittel, PhD

Date Defended: July 1<sup>st</sup>, 2013

The Thesis Committee for KYLE LAMAR MILLER

certifies that this is the approved version of the following thesis:

SPLICING ERROR IN *GATA1* AFFECTS ERYTHROPOIESIS IN THE *XPNA* MOUSE  
(X-LINKED PRE- AND NEONATAL ANEMIA) WITH SUGGESTION OF A NOVEL  
COMPENSATORY ERYTHROID TRANSCRIPTION FACTOR

---

Chairperson Theresa I. Shireman, PhD, RPh

Date approved: July 1<sup>st</sup>, 2013

## Abstract

A novel mutant mouse called X-linked pre- and neonatal anemia (gene symbol, *Xpna*) results in a transient, neonatal anemia which is resolved by 3 weeks of age in *Xpna* females. Adult *Xpna* females exhibit hypoplastic bone marrow with red cell aplasia and splenomegaly showing extramedullary erythropoiesis and megakaryocytosis. We identified a splicing defect derived from a single nucleotide change 5 base pairs downstream of Exon 1 in *Gata1*. The *Xpna* *Gata1* gene produces a transcript, which includes alternative Exon 1Eb/c, known not to induce erythropoiesis. X-chromosome inactivation leads to two populations of hematopoietic cells in *Xpna* females, one of which expresses *Xpna* *Gata1* mRNA. An X-chromosome-associated erythroid genetic marker (*Pgk1*, phosphoglycerate kinase-1) indicates reticulocytes are derived from erythropoietic cells expressing the *Xpna* *Gata1* transcript. These data strongly suggest compensatory gene expression allowing for the generation of erythrocytes despite the lack of GATA-1 production. The *Xpna* mouse could, therefore, lead to the identification of novel erythropoietic factors.

## **Acknowledgements**

First and foremost I would like to thank my mentor Dr. Robert White, who has been an excellent mentor and source of support throughout my three years working with him. You have had a major impact on my scientific development, and for that I am thankful. I am honored to count you as a mentor.

I also would like to thank Dr. Flavia Costa, Dr. Laura Shannon, and Dr. Michael Silvey, who I have had the privilege of working with this year. I am appreciative to you for the help and support in the lab this year. I have no doubt that the projects would not have made as much progress this year without your help.

## Table of Contents

<b>Abstract.....</b>	<b>iii</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>v</b>
<b>List of Tables and Figures.....</b>	<b>vi</b>
<b>Introduction.....</b>	<b>1</b>
Background .....	1
Study Objective.....	2
<b>Materials and Methods.....</b>	<b>2</b>
Mice .....	2
Genotyping.....	3
Hematology.....	3
Histology.....	3
Genetic Mapping.....	4
Mutation Analysis .....	4
Splice Site Analysis .....	5
RNA Extraction .....	5
5'-RACE PCR and Cloning.....	5
RT-PCR.....	6
Real-time Quantitative RT-PCR (qRT-PCR) .....	7
<b>Results .....</b>	<b>9</b>
Inheritance of <i>Xpna</i> Locus .....	9
Hematological Presentation .....	10
Spleen, Liver, and Bone Marrow Histology .....	11
Genetic Mapping.....	13
Mutation Analysis .....	13
5'-RACE PCR.....	16
Absence of Clonal Selection.....	17
RT-PCR Analysis.....	18
<b>Discussion.....</b>	<b>22</b>
<b>Conclusions.....</b>	<b>24</b>
<b>References .....</b>	<b>25</b>

## List of Tables and Figures

Table 1: Primers used for <i>Xpna</i> genotyping .....	3
Table 2: Primers used for 5'-RACE-PCR .....	6
Table 3: Primers used for RT-PCR.....	7
Table 4: Primers used for qRT-PCR .....	8
Figure 1: Phenotype of anemic <i>Xpna</i> pup.....	9
Table 5: Hematologic Data of 2 day old and 3 week old <i>Xpna</i> and Wild type Mice.....	10
Figure 2: Histology of <i>Xpna</i> mice .....	12
Figure 3: Single nucleotide change in <i>Xpna</i> mutation .....	14
Figure 4: <i>Xpna</i> Splice Site analysis .....	15
Figure 5: 5'-RACE PCR.....	17
Figure 6: RT-PCR.....	19
Figure 7: Expression of <i>Gata1</i> in <i>Xpna</i> and wild type mice .....	20
Figure 8: Expression of <i>Gata2</i> in <i>Xpna</i> and wild type mice .....	21

# Introduction

## Background

GATA-1 is an essential transcription factor that regulates erythroid differentiation by targeting specific genes.<sup>1,2</sup> The *Gata1* gene is composed of multiple untranslated first exons and five other common exons.<sup>3</sup> Three functional domains have been described for the GATA-1 protein: an N-terminal domain reported to act as a transcriptional activation domain and two zinc finger domains, one for binding to DNA and another to stabilize DNA binding.<sup>4</sup> GATA1 recognizes the GATA consensus motif, (T/A)GATA(A/G)<sup>5</sup>, and is known to interact with many transcribing factors such as FOG-1, LMO2, KLF1, TAL-1 and PU.1.<sup>1,6-13</sup>

Gene ablation studies of *Gata1* in mice revealed its essential role in erythroid cell development.<sup>14-16</sup> Embryos lacking *Gata1* die at approximately embryonic day 11.5 due to arrested maturation of primitive erythroid cells.<sup>14</sup> Heterozygous *Gata1* knockout females survive after experiencing a transient, neonatal anemia.<sup>14</sup>

Various strategies for gene targeting of *Gata1* in mice have generated different hematological phenotypes. Mouse *Gata1* mutations with inducible knockouts of either Exon 1E (the erythroid first exon) alone or Exons 2 through 6, using the *cre-lox* system, provided insight on the role of the different coding regions of *Gata1*.<sup>3,16</sup> Induction of floxed deletion of Exon 1E, which contains the 5' untranslated region (UTR) of the *Gata1* mRNA, at 3 to 4 weeks of life and resulted in dramatic splenomegaly and severe anemia.<sup>3</sup> In this inducible mutant, an alternative Exon (1Eb/c) is utilized in lieu of Exon 1E, resulting in a complete loss of expression of GATA-1 in the enlarged spleens. The anemia resulting from the inducible knockout persists in these mice, suggesting that Exon 1E is important for GATA-1 expression in hematopoietic lineages. Alternatively, induction of a *Gata1* knockout, including floxed Exon 2 through 6 at 8 to 10

weeks, leads to anemia, thrombocytopenia, and decreased spleen weight.<sup>16</sup> The hematological phenotype of these mice was consistent with the loss of GATA-1 expression.

### **Study Objective**

Our objective in this study is to characterize a unique mouse model known as X-linked pre- and neonatal anemia (gene symbol, *Xpna*), which experiences a transient, neonatal anemia. The *Xpna* female mouse displays a severe anemia at birth that gradually regresses and completely resolves at approximately 3 weeks of age. We identified the mutation associated with *Xpna* mice as a single base pair mutation in intron 1 of the *Gata1* gene. This mutation results in a splicing error leading to the expression of a *Gata1* transcript that utilizes Exon 1Eb/c as an alternative first exon. We hypothesize the recovery from anemia in *Xpna* mice is due to an alternative compensatory mechanism allowing for progression of erythropoiesis.

### **Materials and Methods**

#### **Mice**

The *Xpna* mutant was originally generated by *N*-ethyl-*N*-nitrosurea (enu) mutagenesis on the C3H101H strain (<http://www.informatics.jax.org/reference/J:169366>). C3H101H-*Xpna*/+ (Harwell stock number FESA:000520) was a gift from Bruce Cattanach at the Medical Research Council Harwell (Oxfordshire, England). We transferred the *Xpna* mutation to the C3H/HeJ strain by repeated matings of *Xpna*/+ females to C3H/HeJ male mice (obtained from The Jackson Laboratory, stock number 000659) for several generations to produce a congenic line. All mice were maintained at the University of Missouri-Kansas City (UMKC) Lab Animal Research Center with animal usage guidelines and procedures approved by the UMKC Institutional Animal Care and Use Committee.



## Genotyping

Genomic DNA was extracted from mouse tissue using the Qiagen DNeasy Blood and Tissue kit (Qiagen, Valencia, CA) as described previously.<sup>17,18</sup> The *Xpna* mutation was detected by PCR amplification followed by DNA sequence analysis using primers derived from sequences flanking Exon 1 (Table 1). The PCR product was purified with a GE Illustra Spin column (VWR, Brisbane, CA) and sequenced at the University of Kansas Medical Center Biotechnology Support Facility (Kansas City, KS) using an ABI 377 automated sequencer (Applied Biosystems, Foster City, CA).

Genotyping PCR		
Name	Direction	Sequence
<i>Xpna Gata1</i>	F	GGAGGAAAGAGGAGGGAGAA
	R	CCGTTTtaggggCAACCTAT

**Table 1: Primers used for *Xpna* genotyping**

## Hematology

Peripheral blood samples were obtained in hematocrit tubes from neonatal C3H/HeJ-*Xpna*/+ female pups and normal female littermates. Red blood cells were counted with a Coulter Counter (Model Z1; Miami, FL). Hematocrit percentage was assessed using an Adams Hematocrit Reader (Becton, Dickinson and Company, Parsippany, NJ). Hemoglobin concentration of the red cells lysed in Drabkin's solution (Total Hemoglobin Kit, Sigma Diagnostics, St. Louis, MO) was measured in a Perkin-Elmer Lambda 40 spectrophotometer (Norwalk, CT) at a wavelength of 530 nm.

## Histology

Necropsy, histology and pathology analysis of tissues from *Xpna* mice and wild type littermates at 3 months and 6 months of age was completed by IDEXX Laboratories (Columbia,

MO). Preparation and analysis of bone marrow smears with Wright's-Giemsa stain were completed using techniques as described previously.<sup>19</sup> Spleen and liver samples were fixed, sectioned, and stained with hematoxylin and eosin.

### **Genetic Mapping**

To determine the relative chromosomal location of the *Xpna* mutation, C3H/HeJ-*Xpna*/+ female mice were mated to the CAST/Ei strain (from The Jackson Laboratory, JAX stock number 000664). A backcross of the (C3H/HeJ X CAST/Ei) F<sub>1</sub>-*Xpna*/+ female progeny (identified by pallor, low hematocrit and/or *Gata1* DNA sequencing of the mutation in intron 1) to C3H/HeJ-wild type male mice was used to generate *Xpna*/+ and +/+ offspring. Female mice were the only animals analyzed since *Xpna* males are never born. The *Xpna*/+ mice were scored and genomic DNA was extracted from spleens of the backcross mice using the Super Quik Gene DNA extraction kit (Analytical Genetic Testing Center, Denver CO). These DNA samples were genotyped with primer pairs (Research Genetics, Huntsville, AL) for a series of simple sequence length polymorphisms (CA repeats) as previously described.<sup>20</sup> A total of 13 genetic markers including 12 dinucleotide CA-repeat microsatellite markers and a novel CA repeat identified in GenBank (NT 000086) at position nt4858 in the genomic sequence of *Glod5* (glyoxalase domain containing 5).

### **Mutation Analysis**

Genomic DNAs for mouse strains were obtained from The Jackson Laboratory and the 101H mouse strain sample for DNA extraction was a gift from the Medical Research Council Harwell. Genomic PCR was performed using primers derived from flanking sequence 50 bp upstream and downstream of the 6 exons in *Gata1* using the Fail Safe PCR kit (Epicentre Technologies, Madison, WI). The primers were chosen using the Primer 3 program

([frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)) from the sequence of *Gata1* (GenBank Accession NC\_000086). Analytical gels were 3% GenePure 3:1 agarose (ISCB, Kaysville, UT). PCR products were isolated using CENTRI-SEP Spin Columns (Princeton Separations Inc., Adelphia, NJ) prior to sequencing by the KU Medical Center Biotechnology Support Facility.

### **Splice Site Analysis**

To analyze the status of splice site donors, a perl script was written to extract 189,185 donor sites and 189,092 acceptor sites from the annotated mouse genome. A pseudo-Gibbs motif sampler was run to build the 10-mer donor motif model, and draw a motif logo using BiLogo Plotter (BiLogo Plotter. <http://bipad.cmh.edu/bilogo.html>).<sup>21</sup> Scoring of a potential donor binding site was performed as previously described.<sup>22</sup>

### **RNA Extraction**

RNA extraction from reticulocytes, liver, spleen, fetal liver, and whole embryos was performed using the RNeasy Midi Kit (Qiagen, Valencia, CA). RNA from peripheral blood reticulocytes were obtained from phenylhydrazine-treated mice prepared as previously described.<sup>23</sup>

### **5'-RACE PCR and Cloning**

A 5'-RACE System for Rapid Amplification of cDNA Ends kit (Invitrogen, Carlsbad, CA) was used according to manufacturer's instruction using primers described in Table 2. A dilution of the RACE PCR product was used as the template for nested amplification using a primer that was modified to include *EcoR1* recognition sites. This PCR product was purified with GE Illustra Spin columns (VWR, Brisbane, CA), cut with *EcoR1*, and re-purified with the columns. The cut DNA was ligated into pBlueScript plasmid (Agilent, La Jolla, CA) using T4 DNA Ligase (Promega, Madison, WI) and used to transform competent DH5 $\alpha$  bacterial cells

(Invitrogen, Carlsbad, CA). Plasmid DNA was extracted using the CTAB method.<sup>24</sup> Positive clones were sequenced at the University of Kansas Medical Center Biotechnology Support Facility.

<b>5'-RACE-PCR</b>		
Name	Direction	Sequence
1 <sup>st</sup> Step	R	GCTCTTCCCTTCCTGGTCTT
2 <sup>nd</sup> Step	R	CAGGAATTCCCTCCATACTG
3 <sup>rd</sup> Step	R	GTAGTAGGCCAGTGCTGATG

**Table 2: Primers used for 5'-RACE-PCR**

### **RT-PCR**

*Gata1* mRNA expression was analyzed by RT-PCR amplification using the SuperScript One-Step RT-PCR System for Long Templates (Invitrogen, Carlsbad, CA) using primers derived from the cDNA sequence for detection of *Gata1* mRNA as described in supplemental Table 3. *Pgk1* RT-PCR, using reticulocyte RNA as a template, was performed with primers chosen from the sequence of *Pgk1* (GenBank Accession NT\_039711; Table 1). *Pgk1* RT-PCR products were purified with GE Illustra Spin columns (VWR, Brisbane, CA) and sequenced at the University of Kansas Medical Center Biotechnology Support Facility.

<b>RT-PCR</b>		
Name	Direction	Sequence
<i>Gata1</i> Exon 1-3	F	AATCCTCTGCATCAACAAGC
	R	CATAGGCCCCAGCTAGCATAA
<i>Gata1</i> Exon 2-3	F	CTGGTGCCTCACCATCAGA
	R	CCGTCTTCAAGGTGTCCAAG
<i>Gata1</i> Exon 1E-3	F	ACTCGTCATACCACTAAGGTGGCTG
	R	CACAGTTGAGGCAGGGTAGAG
<i>Gata1</i> Exon 1Eb/c-3	F	GAAGTTGGGGAGCACAGAAG
	R	GCTCTTCCCTTCCTGGTCTT
<i>Pgkl</i>	F	CACGGTGTTGCCAAAATGT
	R	GAAAGCGGAGGTTTTCCAG

**Table 3: Primers used for RT-PCR**

#### **Real-time Quantitative RT-PCR (qRT-PCR)**

cDNA was synthesized using the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA). qPCR analysis was performed with SYBR Green dye using MiniOpticon or CFX96 instruments (Bio-Rad, Hercules, CA). Expression of embryonic and adult  $\beta$ -globin was calculated using the relative quantification method, as previously described (Table 4).<sup>25,26</sup> qRT-PCR was also performed for *Gata1*, *Gata1* Exon 1E, *Gata1* Exon 1Eb/c, *Gata2* and *Gapdh* (glyceraldehyde phosphate dehydrogenase) mRNAs (Table 4). Triplicate data sets were generated and results were normalized to murine *Gapdh* or  $\alpha$ -globin genes (internal controls) for the expression data. Data is shown as the mean  $\pm$  the standard deviation of the mean. The Student's *t*-test was used to determine statistical significance at  $P < 0.05$  and  $P < 0.01$ .

qRT-PCR		
Name	Direction	Sequence
Total <i>Gata1</i>	F	CAGAACCGGCCTCTCATCC
	R	TAGTGCATTGGGTGCCTGC
<i>Gata1 IE</i>	F	ACTCGTCATACCACTAAGGTGGCTG
	R	GCCCCTAGACCAGGAAAATC
<i>Gata1 IEb/c</i>	F	GATAGGGATGAAGTTGGGGAGCAC
	R	GCCCCTAGACCAGGAAAATC
<i>Gata2</i>	F	CACCCCGCCGTATTGAATG
	R	CCTGCGAGTCGAGATGGTTG
<i>Gapdh</i>	F	AGGTTGTCTCCTGCGACTTCA
	R	CCAGGAAATGAGCTTGACAAAG
<i>Hprt</i>	F	GCTGGTGAAAAGGACCTCT
	R	CACAGGACTAGAACACCTGC
<i><math>\alpha</math>-globin</i>	F	GATTCTGACAGACTCAGGAAGAAAC
	R	CCTTTCCAGGGCTTCAGCTCCATAT
<i>Ey</i>	F	CAAGCTACATGTGGATCCTGAGAA
	R	TGCCGAAGTGACTAGCCAAA
<i><math>\beta</math>hl</i>	F	AGGCAGCTATCACAAGCATCTG
	R	AACTTGTCAAAGAATCTCTGAGTCCAT
<i><math>\beta</math>maj+<math>\beta</math>min</i>	F	GTGAGCTCCACTGTGACAAGCT
	R	GGTGGCCCAGCACAAATCACGATC

**Table 4: Primers used for qRT-PCR**

## Results

### Inheritance of *Xpna* Locus

*Xpna* mutant pups are anemic at birth and scored by the observation of pallor as seen in Figure 1. The inheritance pattern is that of an X-linked dominant mutation since only *Xpna*/+ females are observed (0 pale male pups out of 129 pale pups born) and mating *Xpna* females with wild type males results in 38% of the offspring exhibiting the anemia phenotype at birth (29 pale pups out of a total of 77 littermates). Male *Xpna* embryos have been noted to die *in utero* during mid-gestation at embryonic age 10.5 days (personal communication, e-mail, Bruce Cattanaach, Medical Research Council Harwell).



**Figure 1: Phenotype of anemic *Xpna* pup;** A normal pup is pictured on the right and compared to an anemic *Xpna*/+ pup shown on the left with skin pallor.

## Hematological Presentation

Blood samples from 2-day old *Xpna* mice compared to +/+ female littermates showed marked disturbances in red blood cell parameters (Table 5). *Xpna* pups showed a significant reduction in red blood cell number, hematocrit, and hemoglobin. In contrast, blood samples from 3-week old *Xpna* mice showed that the severe, neonatal anemia is resolved establishing that the *Xpna* mutation results in a transient, neonatal anemia. There is also a five-fold increase in nucleated red blood cells in the peripheral circulation of *Xpna* pups which is resolved in adult *Xpna* mice (data not shown). Despite the resolution of the severe anemia, adult *Xpna* females show marked splenomegaly with an average increase in spleen weight of 25% in 3 week old mice to an increase of over 200% in 7 month old mice compared to wild type littermates. Adult *Xpna* females also have mild to moderate thrombocytopenia when compared to wild type littermates ( $386 \pm 112 \times 10^3/\mu\text{L}$  compared to  $738 \pm 144 \times 10^3/\mu\text{L}$ ;  $p < 0.01$ ).

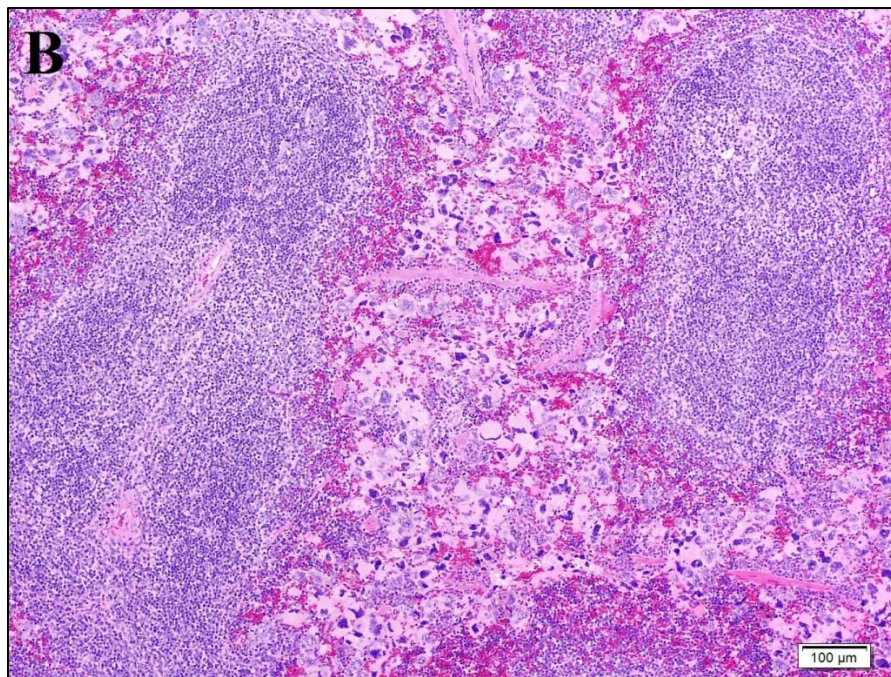
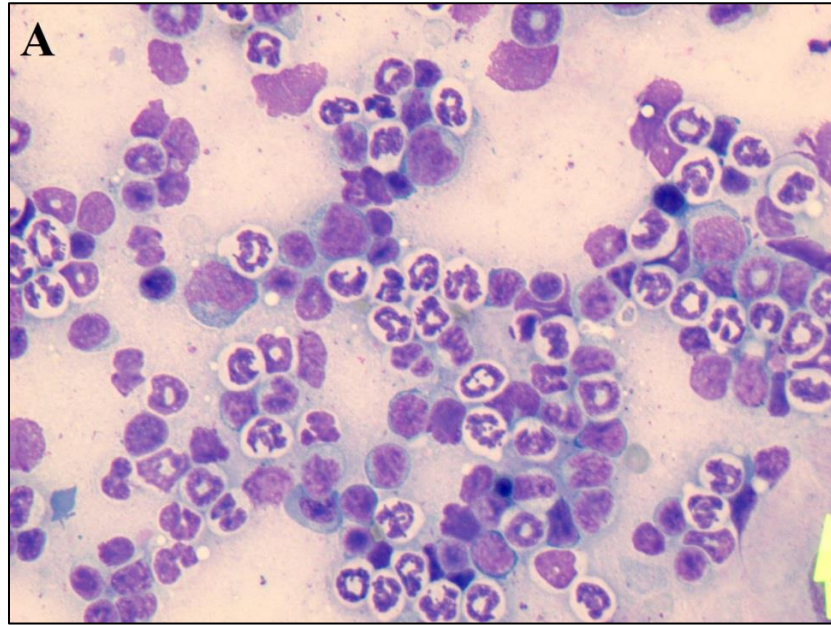
Genotype	RBC # ( $\times 10^{12}/\text{L}$ )	Hct (%)	Hb (g/L)	MCV (fl)
2 day old <i>Xpna</i> /+	$1.48 \pm 0.23^*$	$0.18 \pm 0.04^*$	$67 \pm 11^*$	$124 \pm 37$
2 day old +/+	$3.46 \pm 0.35$	$0.44 \pm 0.05$	$137 \pm 22$	$126 \pm 10$
3-week old <i>Xpna</i> /+	$4.91 \pm 0.48$	$0.39 \pm 0.05$	$117 \pm 13$	$79 \pm 7$
3-week old +/+	$4.52 \pm 0.46$	$0.41 \pm 0.03$	$120 \pm 17$	$77 \pm 9$

**Table 5: Hematologic Data of 2 day old and 3 week old *Xpna* and Wild-type Mice;** Data are mean  $\pm$  standard deviation of the mean. Hgb indicates hemoglobin; Hct, Hematocrit; RBC #, Red Blood Cell Number. \*:Mean value for *Xpna*/+ mice is significantly different from +/+ littermates by *t*-test at  $P < 0.001$ ,  $n=8$ .



### **Spleen, Liver, and Bone Marrow Histology**

Spleen, liver, and bone marrow samples were examined microscopically from adult *Xpna* and +/+ littermates to evaluate for hematological defects. The bone marrow smears of adult *Xpna* females showed predominantly hypoplastic bone marrow with red cell aplasia (Figure 2A). Bone marrow aspirate smears of *Xpna* mice showed varying results diverging from no cytologic abnormalities to marked erythroid hypoplasia. Spleens from *Xpna* mice showed marked megakaryocytosis and mild to moderate hematopoiesis consistently compared to their wild type littermates which showed no megakaryocytosis and zero to mild hematopoiesis (Figure 2B). The histology of *Xpna* liver samples showed no difference compared +/+ littermates.



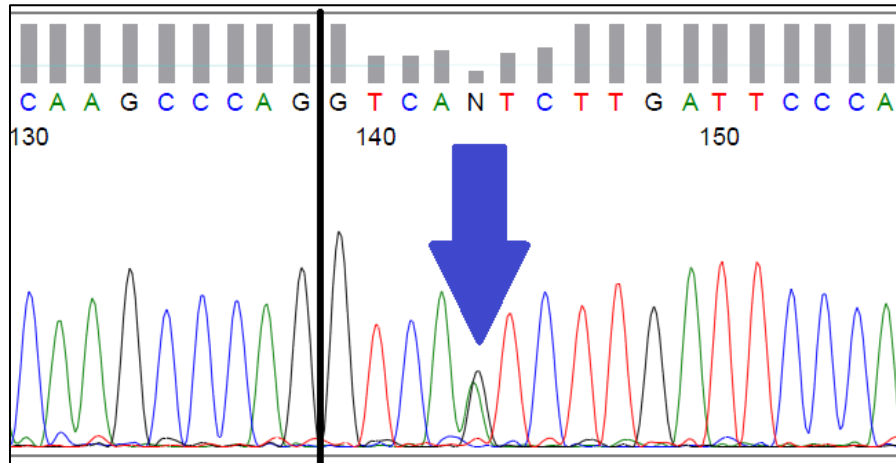
**Figure 2: Histology of *Xpna* mice;** A) Histological appearance of adult *Xpna* bone marrow smear stained with Wright's-Giemsa stain consistent with marked erythroid hypoplasia to pure red cell aplasia. Original magnification x 50. B) Histological appearance of adult *Xpna* spleen stained with hematoxylin and eosin showing mild hematopoiesis and marked megakaryocystosis of the red pulp. Scale bar corresponds to 100  $\mu\text{m}$ .

## Genetic Mapping

Chromosomal localization of the *Xpna* gene was determined by backcross analysis. Exactly 76 backcross mice were genotyped and scored for polymorphic microsatellite genetic markers (dinucleotide repeats) between C3H101H and CAST/Ei. The results of the mapping study indicated that the *Xpna* locus was located on the X chromosome with close linkage to *DXMit26* and *Glod5*, which lies in very close proximity to *Gata1* ([www.informatics.jax.org](http://www.informatics.jax.org)). The similar location of *Xpna* and *Gata1*, along with the resemblance in hematology data between *Xpna* anemia mutant mice and *Gata1* knockout mice, led us to predict that *Gata1* was the most likely gene candidate for the mutation at the *Xpna* locus.<sup>14</sup>

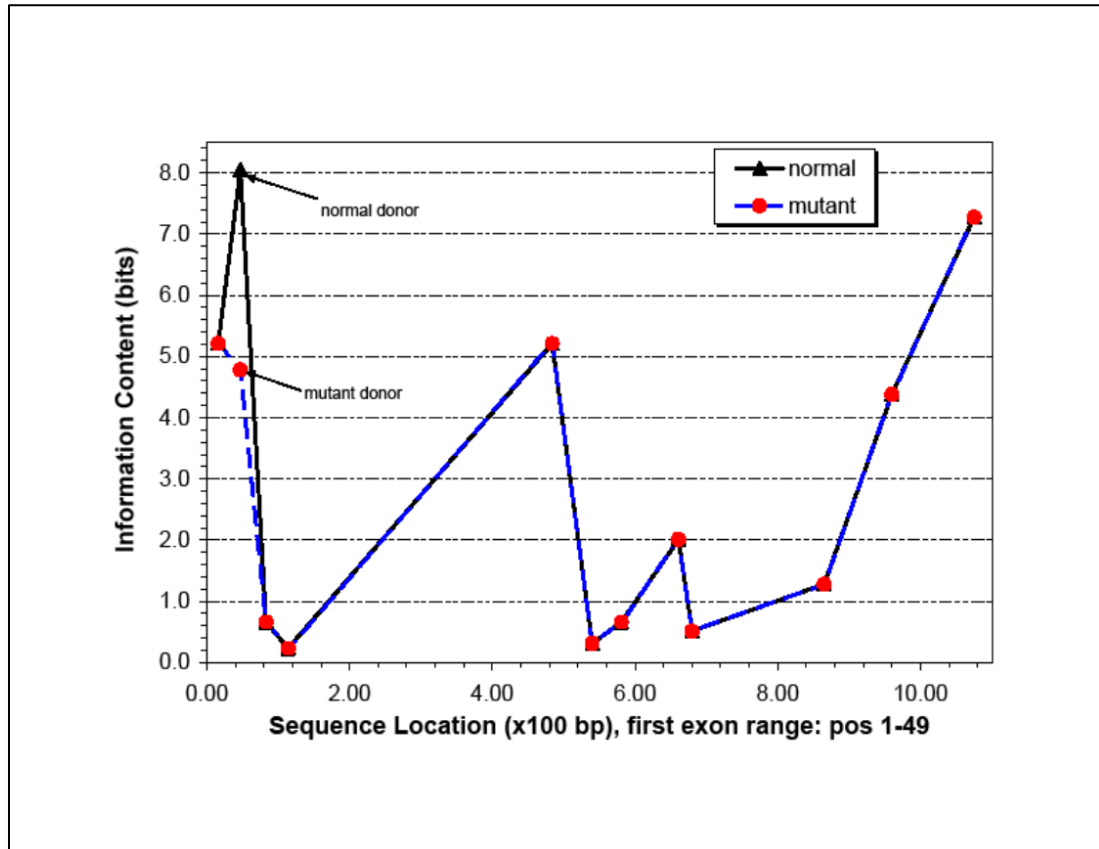
## Mutation Analysis

Comparison of the sequences of the 6 exons of *Gata1* between *Xpna* mice and that of +/+ littermates indicates no change in the coding sequence of *Gata1*. However, there was single nucleotide change seen 5 base pairs downstream of Exon 1E (G to A transition) only seen in mice carrying the *Xpna* locus (Figure 3). To verify that this SNP was unique to *Xpna*, we sequenced this region from 21 different inbred mouse strains including the founder strain for *Xpna* mice (C3H101H), 4 other sublines of the C3H strain (C3H/HeJ, C3HeB/FeJ, C3H/HeSnJ, C3H/HeOuJ) and the 101H parental strain. Analysis of the *Gata1* nucleotide sequence in these different strains demonstrates that mutation is found only in *Xpna* mice.



**Figure 3: Single nucleotide change in *Xpna* mutation;** the chromatogram shows the single nucleotide change (G to A transition; indicated by the blue arrow) that is present 5 base pairs downstream of Exon 1E in the *Gata1* gene of *Xpna* mice. The black bar indicates the division between Exon 1E on the left and intron 1 on the right.

Due to its close proximity to the splicing donor site of exon 1, it was decided to perform an analysis of the donor splice site in *Xpna* mice compared to wild type sequence. A 10-mer donor motif was built to generate a consensus sequence AGGTAAG for donor splice signaling. The donor sequence model was used to scan both normal and mutant sequence and indicated the mutated nucleotide (G->A) has a high impact on the splicing (Figure 4). The normal donor site has the highest information content, indicating the correct binding place. However, the mutated donor decreases information by nearly 2-fold. The relevant binding energy is significantly reduced that may allow its nearby sites to be putative donor sites, causing potential abnormal splicing.

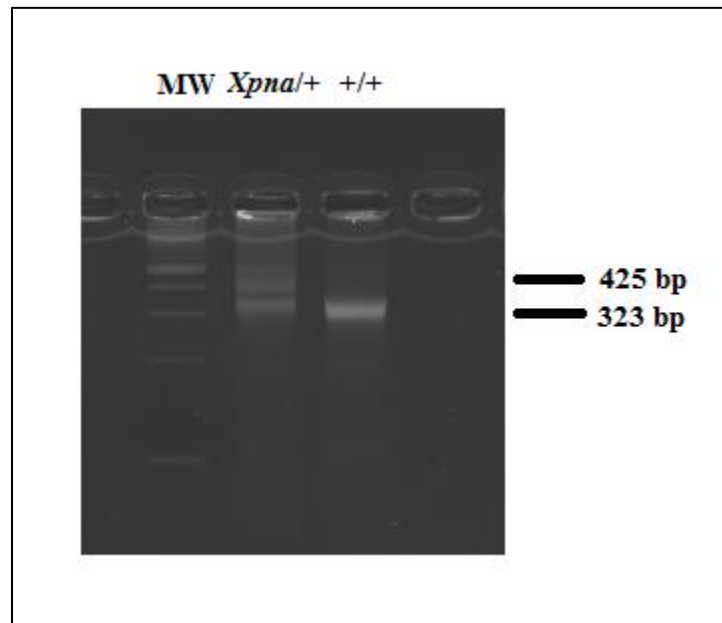


**Figure 4: *Xpna* Splice Site analysis;** the graph shows a decrease in splice site potential, represented by a decrease in information content, due to the single nucleotide change identified in the *Xpna Gata1* allele .

In order to determine if Exon 1E is skipped due to abnormal splicing, RT-PCR of fetal liver RNA from *Xpna*/+ female, *Xpna* male (identified by Y genomic PCR and *Gata1* sequencing), wild type male and female embryos was performed. The results showed that Exon 1E to 3 yielded an appropriate amplicon in all cases except for male *Xpna* embryos whereas products for exon 2 to 3 were seen in all cases including *Xpna* males. This suggested that Exon 1E is not included in the final *Gata1* transcript generated by the *Xpna* locus.

## 5'-RACE PCR

In order to determine what was present at the 5' end of the *Xpna Gata1* mRNA, 5'-RACE PCR was completed using 2 day old liver samples from *Xpna* and +/+ female mice. As shown in Figure 5, the 5'-RACE PCR generated an expected wild type product of 323 bp in *Xpna* and wild type mice. However, a slower migrating and larger size amplicon (425 bp) was detected only in *Xpna* samples. Sequence analysis indicates the wild type band sequence was consistent for both Exon 1E and Exon 2 while a novel DNA sequence upstream of Exon 2 was identified in the *Xpna*-specific band. A BLAST search with the sequence revealed that the upstream sequence seen in *Xpna*-specific band is Exon 1Eb/c. Therefore, the result of the *Xpna* mutation (G to A transition) is a splicing mutation in which Exon 1E is not utilized but instead the alternative Exon 1Eb/c is utilized. The significance of this finding is that *Gata1* mRNA transcripts containing Exon 1Eb/c have been previously reported to be unable to produce full-length GATA1 protein *in vivo*, possibly due to inefficient translation, resulting in impaired erythropoiesis.<sup>3,27</sup>



**Figure 5: 5'-RACE PCR;** 5'-RACE analysis of *Gata1* mRNA shows the presence of an *Xpna* specific transcript that is 425 bp in size. This band is larger and slower migrating than the wild type *Gata1* band which is 323 bp in size.

#### **Absence of Clonal Selection**

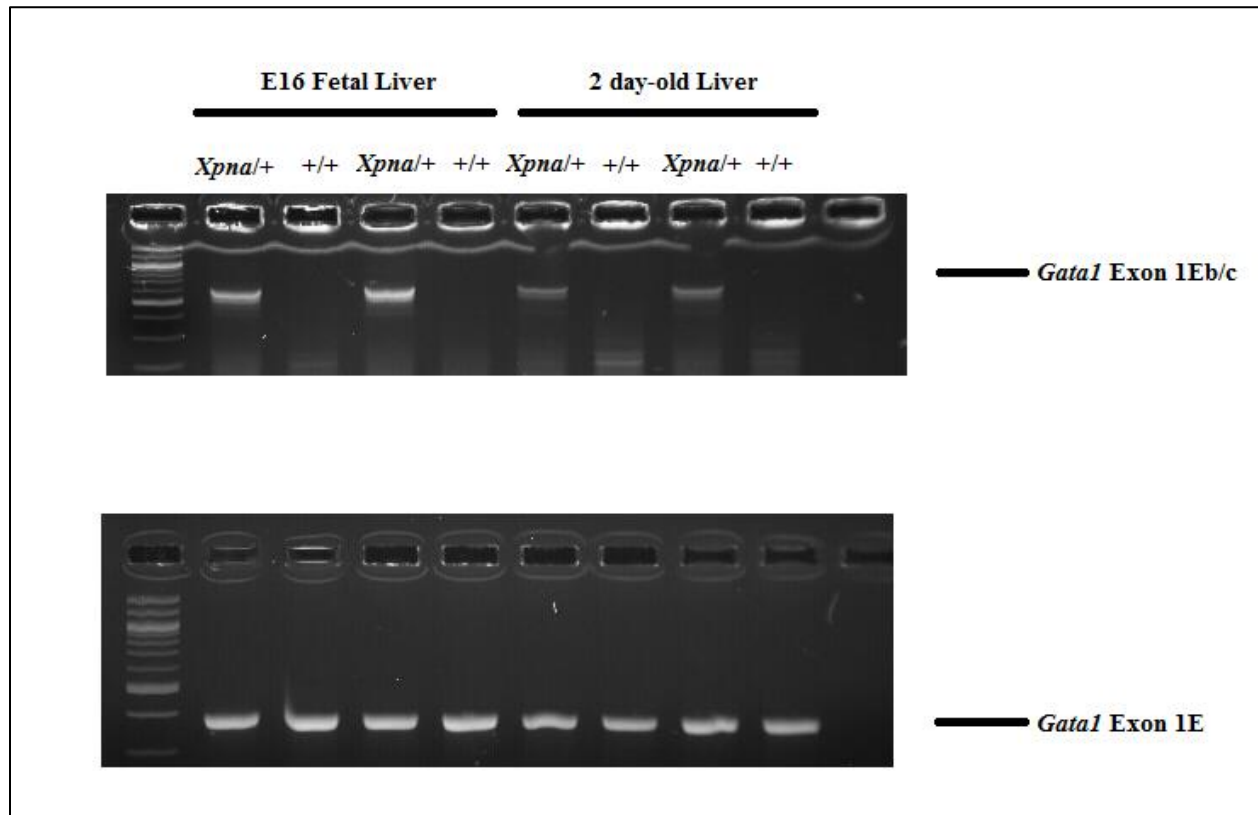
A hypothesis was generated to explain the transitional nature of the *Xpna* anemia in which *in vivo* selection of progenitor cells with an active normal allele is responsible for the recovery from the severe neonatal anemia in *Xpna* females. In order to test the hypothesis we identified an X chromosome, red blood cell genetic marker which differs between C3H101H *Xpna*/+ mice and a polymorphic strain CAST/Ei (C>T at position 292 in exon 3 coding region of *Pgk1* mRNA; rs29095411 SNP at position 106194380nt on the mouse X chromosome). These mice were mated to generate (C3H101H X CAST/Ei) $F_1$  *Xpna*/+ and ++ female mice carrying a X chromosome of C3H101H origin and a X chromosome from the CAST/Ei strain. Reticulocytosis was induced and reticulocyte RNA was isolated from peripheral blood for RT-PCR analysis. Sequencing of the *Pgk1* amplicons was utilized to ascertain if clonal selection of

erythroid cells with an active CAST/Ei-derived X chromosome was observed in *Xpna* mice. Sequencing analysis from 3 (C3H101H X CAST/Ei)F<sub>1</sub> wild type +/+ females and 3 (C3H101H X CAST/Ei)F<sub>1</sub> *Xpna*/+ females revealed an identical pattern of expression for both forms of *Pgk1* mRNA. This indicates the absence of clonal selection in *Xpna* mice and strongly suggests that red blood cells are produced from progenitor cells that can only express the *Xpna Gata1* mutant.

### **RT-PCR Analysis**

RT-PCR and qRT-PCR analysis was utilized to evaluate the expression of *Gata1* Exon 1E and Exon 1Eb/c in *Xpna* mice. This was completed using RNA extractions from day E16 fetal liver, 2 day-old liver and adult spleen. RT-PCR analysis as shown in Figure 6A and Figure 6B confirms that *Gata1* mRNA containing Exon 1Eb/c is only expressed in the fetal and 2 day-old liver of *Xpna* mice, whereas transcripts containing Exon 1E are expressed in both *Xpna* and wild type mice. qRT-PCR analysis revealed that E16 fetal *Xpna* liver had 1.4 fold increase in total *Gata1* mRNA levels, a 0.45 fold decrease in *Gata1* Exon 1E mRNA levels and a 5.5 fold increase in *Gata1* Exon 1Eb/c mRNA levels (Figure 7A). 2 day-old *Xpna* liver was found to have a 5 fold increase in total *Gata1* mRNA levels, a 1.8 fold increase in *Gata1* Exon 1E mRNA levels and a 65 fold increase in *Gata1* Exon 1Eb/c mRNA levels (Figure 7B). Adult *Xpna* spleen was found to have a 4.3 fold increase in total *Gata1* mRNA levels, a 1.5 fold increase in *Gata1* Exon 1E mRNA levels and a 146 fold increase in *Gata1* Exon 1Eb/c mRNA levels (Figure 7C). The high levels of *Gata1* Exon 1Eb/c expression further suggests that clonal selection is not occurring and supports our hypothesis that red blood cells are being produced from progenitor cells that express the *Xpna Gata1* mutant.



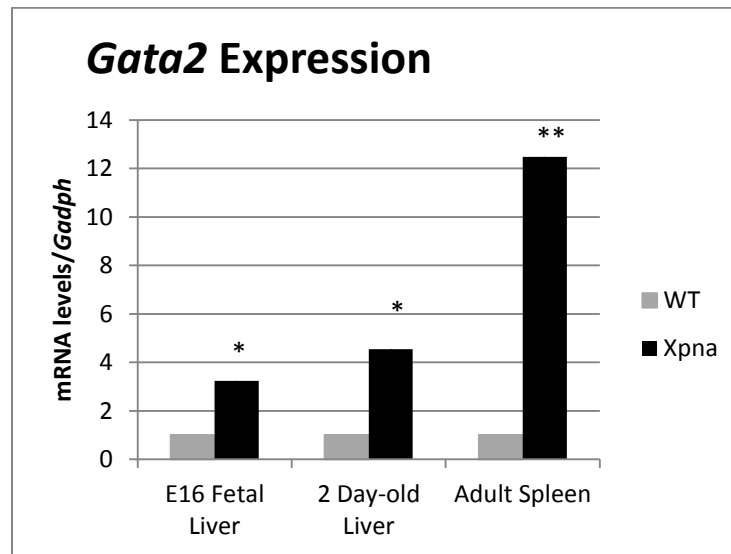


**Figure 6: RT-PCR;** analysis of *Gata1* mRNA from E16 fetal liver and 2 day-old liver confirms that *Gata1* Exon 1Eb/c is only expressed in *Xpna*, whereas *Gata1* Exon 1E is expressed in wild type and *Xpna*/+ mice.



**Figure 7: Expression of *Gata1* in *Xpna* and wild type mice;** A) expression of total *Gata1* mRNA and *Gata1* Exon 1Eb/c is significantly elevated and *Gata1* Exon 1E mRNA is significantly reduced in E16 *Xpna* fetal liver (n=4). B) expression of total *Gata1* mRNA is significantly increased between 2 Day-old wild type and *Xpna* liver (n=6). C) expression of *Gata1* mRNA is not significantly different between adult wild type and *Xpna* spleen (n=4).

To understand if known genes represent the compensatory network allowing progenitor cells expressing the *Xpna Gata1* mutation to generate red blood cells, we examined the expression by qRT-PCR of mouse hemoglobins mRNA. In fetal liver, 2 day-old liver and adult *Xpna* spleen, there was no detection of changes in hemoglobin mRNA expression in comparison to wild type littermates showing that a reversion to primitive erythropoiesis instead of definitive erythropoiesis in these mutant mice is not the compensatory mechanism for their recovery (data not shown). Expression of *Gata2* mRNA was found to be significantly elevated in *Xpna* mice at all three age points when compared to wild type littermates. There was a 3.2 fold increase in mRNA levels in fetal liver, a 4.6 fold increase in mRNA levels in 2 day-old liver, and a 12.5 fold increase in mRNA levels in adult spleen (Figure 8). Elevated expression of *Gata2* mRNA was expected in *Xpna* mice due to the presence of megakaryocytosis.



**Figure 8: Expression of *Gata2* in *Xpna* and wild type mice;** expression of *Gata2* mRNA is significantly increased in E16 fetal liver, 2 day-old liver, and adult *Xpna* spleen compared to wild type mice.

## Discussion

In this study we have characterized and identified the *Xpna* hereditary anemia mutation. We found by positional cloning that the *Xpna* locus mapped near the *Gata1* gene on the X chromosome. Furthermore, similarities between the *Xpna* hematological phenotype and that of mice heterozygous for the GATA-1-Null allele, which are anemic at birth and recover during the neonatal period, lead us to predict that *Xpna* was a *Gata1* mutation. We identified a SNP in the intron 1 region of the *Gata1* gene in *Xpna* mice, which was predicted to alter mRNA splicing. This was confirmed with 5'-RACE PCR, in which we identified that the *Xpna* mutation results in the utilization of the alternative first exon, Exon 1E b/c, in the mRNA transcript instead of Exon 1E. The anemia that is a result of the splicing error is consistent with a previous study in male mice with a *Gata1* Exon 1E knockout allele (*Gata1*<sup>IECKO/Y</sup>), in which the *Gata1* transcript containing Exon 1Eb/c was shown to be unable to produce a full length GATA1 protein therefore impairing erythropoiesis.<sup>3</sup>

We set out in this study to identify possible mechanisms in which the *Xpna* mice recover from anemia during the neonatal period. The recovery from anemia seen in the heterozygous GATA-1-Null mouse, which also has transitional neonatal anemia, has been hypothesized to be due to the *in vivo* selection of progenitor clones with an active normal allele.<sup>14,15</sup> In order to test if clonal selection was present in adult *Xpna* mice we utilized the X chromosome associated red blood cell marker *Pgk1*. Data from our study found that reticulocytes were present from two distinct cell populations in the mice based upon X chromosome inactivation. This suggests that clonal selection is not occurring and therefore not responsible for the transitional anemia in *Xpna* mice.

The alternative shorter isoform of GATA-1, GATA-1s, which lacks the N-terminal activation domain, was also hypothesized to be a potential mechanism for the transitional anemia. Overexpression of GATA-1s in male hemizygous *Gata1*<sup>ΔN</sup> mice, which has an in-frame deletion of codons for amino acids 3 to 63 in Exon 2 of Gata1, was shown to support normal adult megakaryopoiesis, platelet formation, and erythropoiesis with no disturbances to hematological parameters or spleen structure.<sup>28</sup> Expression of GATA-1s at levels comparable that were comparable to endogenous GATA-1 expression using the *Gata1*<sup>ΔNT</sup> mutant, which has a deletion of N-terminal amino acids 1 through 83, was shown to not be sufficient to sustain definitive erythropoiesis.<sup>29</sup> This suggests that GATA-1s is only able to sustain erythropoiesis when expression is higher than that of endogenous GATA-1.

Exon 1Eb/c containing *Gata1* transcripts were shown in adult *Gata1*<sup>IECKO/Y</sup> mice to predominantly result in the expression of GATA-1s at low levels, due to inefficient translation.<sup>3</sup> Additionally, immunohistochemical analysis of the spleens in these mice showed that full-length GATA-1 was not detectable, which suggests that Exon 1Eb/c does not have the ability to produce full-length GATA-1 protein.<sup>3</sup>

The lack of evidence of clonal selection suggests that based off of X chromosome inactivation, there are two erythroid lineages present in adult *Xpna* mice. Our data from the *Pgkl* study suggests that erythroid progenitor cells which have an active X chromosome containing the *Xpna Gata1* allele are able to produce mature red blood cells despite only expressing the Exon 1Eb/c form of *Gata1* mRNA. When considering the results of the *Gata1*<sup>IECKO/Y</sup> study, we speculate that these erythroid progenitor cells in *Xpna* mice are able to produce mature red blood cells without the expression of full-length GATA-1 protein. Furthermore, the inefficient translation of Exon 1Eb/c containing *Gata1* transcripts suggests that erythropoiesis is occurring

in the cell lineages with minimal expression of GATA-1s. We hypothesize that erythropoiesis is occurring in these cell lineages that only express the *Xpna Gata1* allele due to a compensatory erythropoietic factor. We speculate that this compensatory mechanism is developmentally expressed and is responsible for the recovery from anemia in the neonatal period. We further speculate that this compensatory factor does not impact megakaryopoiesis. The spleens in adult *Xpna* had marked megakaryocytosis and elevated levels of *Gata2* mRNA. These findings are consistent with the results of GATA-2 overexpression in GATA-1s knockin mice, which exhibited aberrant megakaryopoiesis.<sup>30</sup> Further studies of the *Xpna* mice will be needed will be needed to identify compensatory factor.

## Conclusions

In conclusion we have characterized the phenotype and genotype of the *Xpna* mutation. *Xpna*/+ mice are characterized by a transient, neonatal anemia due to splicing mutation of the *Gata1* gene. The *Xpna* mouse mutant acts as a model for studying *Gata1* alternative first exon 1Eb/c. The mechanism for recovery from the severe, neonatal anemia is not known, but our study suggests that a compensatory erythropoietic factor is the likely mechanism. Further research is warranted to identify novel erythropoietic factor.

## References

1. Cantor AB, Orkin SH. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene*. 2002;21(21):3368-3376.
2. Ferreira R, Ohneda K, Yamamoto M, Philipsen S. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol Cell Biol*. 2005;25(4):1215-1227.
3. Kobayashi E, Shimizu R, Kikuchi Y, Takahashi S, Yamamoto M. Loss of the Gata1 gene IE exon leads to variant transcript expression and the production of a GATA1 protein lacking the N-terminal domain. *J Biol Chem*. 2010;285(1):773-783.
4. Martin DI, Orkin SH. Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes Dev*. 1990;4(11):1886-1898.
5. Merika M, Orkin SH. DNA-binding specificity of GATA family transcription factors. *Mol Cell Biol*. 1993;13(7):3999-4010.
6. Tsang AP, Visvader JE, Turner CA, et al. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*. 1997;90(1):109-119.
7. Wadman IA, Osada H, Grutz GG, et al. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J*. 1997;16(11):3145-3157.
8. Blobel GA, Nakajima T, Eckner R, Montminy M, Orkin SH. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc Natl Acad Sci U S A*. 1998;95(5):2061-2066.
9. Zhang P, Behre G, Pan J, et al. Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc Natl Acad Sci U S A*. 1999;96(15):8705-8710.

10. Rekhtman N, Radparvar F, Evans T, Skoultschi AI. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev.* 1999;13(11):1398-1411.
11. Nerlov C, Querfurth E, Kulesa H, Graf T. GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood.* 2000;95(8):2543-2551.
12. Zhang P, Zhang X, Iwama A, et al. PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood.* 2000;96(8):2641-2648.
13. Cantor AB, Katz SG, Orkin SH. Distinct domains of the GATA-1 cofactor FOG-1 differentially influence erythroid versus megakaryocytic maturation. *Mol Cell Biol.* 2002;22(12):4268-4279.
14. Fujiwara Y, Browne CP, Cunniff K, Goff SC, Orkin SH. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci U S A.* 1996;93(22):12355-12358.
15. McDevitt MA, Shivdasani RA, Fujiwara Y, Yang H, Orkin SH. A "knockdown" mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. *Proc Natl Acad Sci U S A.* 1997;94(13):6781-6785.
16. Gutierrez L, Tsukamoto S, Suzuki M, et al. Ablation of Gata1 in adult mice results in aplastic crisis, revealing its essential role in steady-state and stress erythropoiesis. *Blood.* 2008;111(8):4375-4385.
17. White RA, McNulty S, Roman S, et al. Chromosomal localization, hematological characterization, and iron metabolism of the hereditary erythroblastic anemia (*hea*) mutant



- mouse. *Blood* 2004;104(5): 1511-1518.
18. White RA, McNulty SG, Nsumu NN, Boydston LA, Brewer BP, Shimizu K. Positional Cloning of the *Ttc7* Gene Required for Normal Iron Homeostasis and Mutated in *hea* and *fsn* Anemia Mice. *Genomics* 2005;85(3): 330-337.
  19. Provencher A, Bolliger AP. Cytologic evaluation of bone marrow in rats: indications, methods, and normal morphology. *Vet. Clin. Pathol.* 2004;33(2):58-67.  
2000;25(4):440-443.
  20. Dietrich W, Katz H, Lincoln SE, et al. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* 1992;131(2):423-447.
  21. Bi C. Comparison of optimization techniques for sequence pattern discovery by maximum-likelihood. *Pattern Recognition Letters* 2010; 31(14): 2147-2160.
  22. Bi C. Memetic algorithms for de novo motif-finding in biomedical sequences. *Artificial Intelligence in Medicine* 2012; 56 (1): 1-17.
  23. White RA, Birkenmeier CS, Lux SE and Barker JE. Ankyrin and the hemolytic anemia mutation, *nb*, map to mouse Chromosome 8: Presence of the *nb* allele is associated with a truncated erythroid ankyrin, *Proc. Natl. Acad. Sci. USA* 1989;83 (4) 3117-3121.
  24. Del Sal G, Manfioletti G, Schneider C. The CTAB DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. *Biotechniques* 1989; 7(5):514-520.
  25. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time

- quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* (2001);25(4): 402-408.
26. Pfaffl MW. A new mathematical model for relative quantification in Real-time PCR. *Nucleic Acids Res.* 2001;29(9): 2002-2007.
  27. Hollanda LM, Lima CS, Cunha AF, et al. An inherited mutation leading to production of only the short isoform of GATA-1 is associated with impaired erythropoiesis. *Nat Genet.* 2006;38(7):807-812.
  28. Li Z, Godinho FJ, Klusmann JH, Garriga-Canut M, Yu C, Orkin SH. Developmental stage-selective effect of somatically mutated leukemogenic transcription factor GATA1. *Nat Genet.* 2005;37(6):613-619.
  29. Shimizu R, Takahashi S, Ohneda K, Engel JD, Yamamoto M. In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *EMBO J.* 2001;20(18):5250-5260.
  30. Huang Z, Dore LC, Li Z, et al. GATA-2 reinforces megakaryocyte development in the absence of GATA-1. *Mol Cell Biol.* 2009;29(18):5168-5180.